

THE DOLEZAL EFFECT

THE DOLEZAL EFFECT: AN INTERACTION EFFECT AND 2 OTHER STUDIES

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirement for the Degree of

Master of Arts

by

Michael David Creim

December 2018

THE DOLEZAL EFFECT

© 2018 Michael David Creim

ABSTRACT

The psychological literature is replete with examples of bias against underrepresented minority groups. However, there is a comparative lack of research on promotion decisions, which the studies here examine in promote or fire scenarios. There is also an emerging small body of literature which suggests that the traditional biases may reverse under certain circumstances such as high prestige/status and less cognitive load, suggesting that formal promotion decisions could be among the exceptions. Study 1 finds a main effect of a pro-Black/anti-White bias in a promote or fire scenario. Study 2 replicates this effect finding that White profiles were less likely to be promoted than a Black profile condition as well as a no race control condition. Study 3 while replicating this main effect found it to be driven by an interaction effect of race and racial stereotypicality where the White employee who diverged from the White stereotype was least likely to be promoted. This paper labels this interaction effect, an apparent aversion to White persons deviating from cultural stereotypes, the Dolezal Effect.

THE DOLEZAL EFFECT

Biographical Sketch

Michael Creim graduated with a Bachelor of Arts degree in Psychology from The University of California, Los Angeles in 2012. He is pursuing his Ph.D. in Developmental Psychology in the Department of Human Development at Cornell University in conjunction with his J.D. at Cornell Law School. His research explores the Psychology of stigma, prejudice, and intergroup differences, as well as sense of purpose in life, and how these topics intersect with the law.

ACKNOWLEDGMENTS

I would like to thank Dr. Anthony Burrow for his guidance and support throughout this project. I would also like to thank all the members of my committee, Dr. Wendy Williams, Dr. Stephen Ceci, and Dr. Valerie Hans. I am also grateful for the numerous undergraduate research assistants who helped with various phases of this project from assembly of materials to data collection.

TABLE OF CONTENTS

	Page
Abstract.....	3
Biographical Sketch.....	4
Acknowledgements.....	5
Table of Contents.....	6
Introduction.....	8
A Gap in the literature.....	9
Exceptions to the Rule of Thumb.....	10
Two Contrasting Predictions.....	12
Stimuli Pilot.....	13
Study 1.....	14
Participants.....	14
Methods.....	15
Results.....	16
Study 2.....	18
Participants.....	18
Methods.....	19
Results.....	20
Study 1 & 2 Combined Meta-Analysis Results.....	22
Study 1 and 2 Discussion.....	22
Study 3.....	24
Participants.....	24
Methods.....	24
Results.....	26

THE DOLEZAL EFFECT

Discussion/Conclusion.....	31
Future Directions/Limitations.....	33
References.....	37

THE DOLEZAL EFFECT

The Dolezal Effect: An Interaction Effect and 2 Other Studies

The overarching theme of work on bias and prejudice is clear. There is bias against minorities, both women in underrepresented fields (Garcia-Retamero & López-Zafra 2006) and ethnic minorities (e.g. Greenwald & Pettigrew 2014; Rosette, Leonardelli, & Phillips 2008).

White persons with implicit biases against ethnic minorities also demonstrate greater explicit biases and behavior (McConnell & Leibold 2001). Defendants with greater Afrocentric features receive harsher criminal sentences (Blair, Judd, & Chapleau 2004). In Georgia Black defendants received harsher sentencing for the same crime, and within race darker skinned individuals also receive harsher sentences (Burch 2015). There are also demonstrations of racial shooter bias, where in shoot/don't shoot simulations, unarmed Black targets are more likely to be shot in error and take longer for participants to decide not to shoot than unarmed White targets, similarly participants are faster to correctly decide to shoot armed Black targets than armed White targets (Correll, Park, Judd, & Wittenbrink 2002).

Aversive Racism Theory provides an important framework for understanding these biases particularly in the context of race (e.g. Dovidio & Gaertner 2000; Dovidio & Gaertner 2004). According to Aversive Racism theory many people explicitly endorse egalitarian views and regard themselves as non-prejudiced, yet nonetheless when faced with ambiguous contexts, where they can ascribe some reason other than race to their decision, will systematically discriminate in favor of White persons. Namely, unlike the bigots of the pre-civil rights era who openly flaunted their racism most people now do not consciously seek to be prejudiced (though unfortunately some still do (e.g. Murphy 2017)). Instead racial bias creeps in at ambiguous decisions and borderline cases where the decisional reasoning can be consciously attributed to something else (Dovidio & Gaertner 2000; Dovidio & Gaertner 2004). In fact, the desire to not

THE DOLEZAL EFFECT

be racist can be a source of anxiety (Mendes & Koslov 2013) and can disrupt White participant's attentional focus (Richeson & Trawalter 2008).

One area where these biases have been well documented is in hiring studies. White resume names are much more likely to receive a call back than black resume names with identical resumes (Bertrand & Mullainathan 2004). In fact, minority job applicants with no criminal record fair equivalently to Whites with a criminal background for entry level jobs (Pager, Bonikowski, & Western 2009). And things have not gotten better over time, as a meta-analysis of 24 studies shows that bias against hiring black employees for entry level (low status jobs) has remained stable for at least a quarter century (Quillian, Pager, Hexel, & Midtbøen 2017).

Many scales and measures have been created to measure the biases which cause these effects. This includes both instruments designed to assess explicit and implicit biases such as the Internal and External Motivation to Respond without Prejudice Scale (Plant & Devine 1998) the Implicit Associations Test (Nosek, Banaji, & Greenwald 2002), and the Color-blind Racial Attitudes Scale (Neville et al 2000).

A Gap in the Literature

Despite the wealth of studies on discrimination in hiring, there are comparatively few attempts to unpack the psychological mechanisms at play in professional advancement or promotion settings. For example, for sexual minorities it is documented that perceived workplace discrimination is linked with lower rates of promotion (Ragins & Cornwell 2001), but the precise decisional mechanisms causing this haven't necessarily been examined in laboratory settings. One of the few studies to examine promotion decisions in lab (Garcia-Retamero & López-Zafra 2006), focused on gender stereotype congruity and leadership positions. Regular

interstitial promotions, for positions between new employee and top-level leadership have been largely ignored.

There is of course documented evidence about the underrepresentation of women and ethnic minorities in advanced positions such as underrepresentation of women and ethnic minority persons who are partners at law firms (e.g. Dinovitzer et al 2014). However, what role promotion decisions play in this underrepresentation and the psychological mechanisms at work in promotion contexts have been underexplored. The inquiry of the studies in this paper begins as an attempt to explore this gap surrounding promotion decisions.

Exceptions to the Rule of Thumb

The general theme of work on prejudice and stigma is well known and highly documented; however, there is an emerging body of literature that offers some nuance and exceptions.

Mendes and Koslov (2013) found that those more biased against minorities were prone to overcorrection when not under cognitive load. In a series of experiments they found that White participants smiled more when interacting with Black confederates than White confederates (and when interacting with confederates with mild make-up based facial disfigurement), as well as demonstrating a preference for minority celebrities, which disappeared when the participants were placed under cognitive load. Similarly, and most apt to the present inquiry Mendes and Koslov (2013) also found that participants with the greatest anti-Black bias on the IAT, when not under threat in a resume selection task were likely to endorse a greater number of medium qualified Black candidates than medium qualified White candidates, but that when under cognitive load/threat they were more likely to endorse a greater number of White candidates (participants low in anti-Black IAT bias were unaffected). Cognitive load is not the only moderator for bias in various contexts. For example, Burrow and Hill (2013) show that greater

THE DOLEZAL EFFECT

sense of purpose in live corresponds to increased comfort with ethnic minorities, with purpose acting as a moderating buffer for anxiety caused by diversity.

Williams and Ceci (2015) found a 2 to 1 hiring advantage for equally qualified female professors in STEM and Axt Ebersol and Nosek (2016) found across 6 studies a bias towards admitting Black students as new members of a campus honors society. Ceci & Williams (2015) contextualizing this with an important boundary condition. The pro women bias did not extend to unqualified candidates and only works in the zone of acceptable candidates, thus the pro-under represented group bias operates in the same space of ambiguity as does aversive racism. An important commonality that Williams and Ceci (2015) and Axt Ebersol and Nosek (2016) share is that simply by virtue of being in consideration for the position applicants had already shown they were incongruent with their negative stereotype. Both professor jobs and honors society membership are high status positions.

Zigerell (2018) conducted a Meta-Analysis using the Time-sharing Experiments for the Social Sciences pool of 17 survey experiments where he found that white participants on average did not exhibit a pro or anti minority bias, while minority participants exhibited a small pro minority bias. Zigerell's study is notable for including 4 workplace/hiring studies and for exposing and noting the possibility of publication bias in the under reporting in the published literature of these sorts of studies.

One of the more curious examples of these exceptions is Mizel (2018). Mizel (2018) found interaction effects in three instances for pretrial decision making where higher implicit racial bias (more pro-White anti-Black) led to worse results for White defendants. Specifically, pro-White implicit racial bias (as assessed by the IAT), caused bail recommendations to be set at higher values for White defendants, target plea sentences recommendations to be longer for the White defendants, and fewer charges were recommended to be dropped for White defendants.

THE DOLEZAL EFFECT

In contrast implicit racial bias did not affect recommendations for Black defendants for their bail determinations, sentence length recommendations, or decisions on how many charges to recommend dropping. It is also worth noting that the sample population Mizel (2018) did overall indicate less of a pro White implicit racial bias on the IAT than the national average. While the other studies that make up this exception group may be interpreted as a positive bias towards the minority group, Mizel's (2018) results instead seem to indicate a willingness to punish Whites for deviating from their positive stereotype, by those most likely to implicitly hold those positive stereotypes about Whites.

Collectively these studies suggest that under certain conditions a pro-minority bias may occur. Specifically, Mendes and Koslov (2013) suggests that the absence of cognitive load is an important factor, thus more formal decisions with greater thought devoted to them may have greater likelihood of demonstrating a pro-minority bias. Mizel (2018) suggests that rather than a pro-minority bias, one may find anti-White biases when Whites deviate from their positive stereotypes. Williams & Ceci (2015) and Axt Ebersol and Nosek (2016) found that underrepresented groups (women in STEM in the case of Williams & Ceci (2015) or Black students in honors societies in the case of Axt Ebersol and Nosek (2016)) may reap benefits in situations where they clearly belie their group's negative stereotype, as to even be considered for a faculty job or honors society requires a minimum level of qualification that runs counter to the negative stereotype. And importantly as Ceci & Williams (2015) indicates, there is a boundary condition where this pro-minority bias does not extend to unqualified individuals. Though, Mendes and Koslov (2013) indicates it may extend to ambiguously qualified individuals.

Two Contrasting Predictions

Consequently, the available literature on bias sets up two contrasting predictions for the impact of race on promotion decisions. On the one hand there is the vast majority of the

THE DOLEZAL EFFECT

literature (Pager, Bonikowski, & Western 2009; Quillian et al. 2017; Dovidio & Gaertner 2004;) which points to the typical biases in hiring decisions against minorities/under-represented groups, and which suggests the same pattern is likely to replicate in a promotion decision context. On the other hand, the special nature of promotion decisions makes them a prime contender to be one of the enumerated exceptions. Under represented persons who are being considered for promotion are by virtue of their present employment already in contradiction with their group's negative stereotype (Williams & Ceci 2015 ; Axt Ebersol & Nosek 2016) like how masculine gay men are at less risk of negative affect (Glick et al. 2007).

Promotion decisions are large decisions for which significant cognitive resources are deployed and generally not undertaken in a hurry while under cognitive load (Mendes & Koslov 2013). Therefore, the list of exceptions, narrow though they are, suggests that promotion decisions may also be a realm of pro-minority bias. Specifically, in a formal promotion decision context, ambiguously qualified underrepresented minorities may receive a boost in consideration since they are clearly inconsistent with their negative group stereotype. Conversely, like the case of the more effeminate men who face greater negative affect for deviating from their positive stereotype of masculinity (Glick et al 2007), ambiguously qualified White employees who deviate from their positive White cultural stereotype may be less likely to receive a promotion.

Stimuli Pilot

Participants:

15 undergraduate students approximately half of which were drawn from our research lab volunteered for our photo stimuli pilot.

Methods:

Studies 1, 2, and 3 would make use of four faces (shoulder height and above) of men shown wearing professional business attire, two of the photos were of White men, two were of

THE DOLEZAL EFFECT

Black men, two were of young men and two were of middle-aged men. Before study 1 we pretested these photos for measures of age and attractiveness.

Results:

The young White photo was perceived on average as having an age of 27.67, with a range from 23 to 32. The young Black photo was perceived on average as having an age of 28.47, with a range from 23 to 34. The middle aged White photo was perceived on average as having the oldest age of 41.06, with a range from 32 to 50. The middle-aged Black photo was perceived on average as having an age of 36, with a range from 29 to 55.

Attractiveness was measured on a 7 point Likert scale ranging from very unattractive to very attractive. Despite the photos being of professional models in professional business attire there was some degree of a floor effect. The range for the middle aged White photo was 1 (very unattractive) to 4 (neither attractive nor unattractive), while the range for the other 3 photos was a similar 1 (very unattractive) to 5 (somewhat attractive). The mean attractiveness for the middle age White photo was 2.13. The mean attractiveness for the middle age Black photo was 3.00. The mean attractiveness for the young White photo was 2.93. The mean attractiveness for the young Black photo was 3.93. The young Black photo had the highest attractiveness, yet was still below the midpoint on the scale.

Study 1

Participants:

Participants for Study 1 consisted of 188 Cornell undergraduates recruited through SONA who participated for the compensation of course credit. Approximately two thirds of participants (123) were female. Of the participants, 81 were White/Caucasian, 47 were Asian/Asian-American, 18 were Black/African-American, 9 were Latino, 15 were multiracial, and 18 identified as other or declined to state. Participants ages ranged from 18 to 27 with over

THE DOLEZAL EFFECT

90% between 18 and 22. Participants were drawn from all 7 of Cornell's undergraduate college and represented a diverse array of family financial backgrounds with incomes ranging from less than \$20,000 to more than \$200,000. The median household income was between \$110,000 and \$140,000, and the modal income was more than \$200,000.

Methods:

Design:

The current study employed a 2 (Stimuli Race: Black vs. White) x 2 (Stimuli Age: young vs. middle aged) design.

Procedure:

Participants came into the lab, where upon a research assistant assigned them to a computer station with an informed consent document on the screen. After providing their informed consent participants were told that we were seeking their input on “a real-life situation” and that we would also be collecting info on their “personality types and learning styles”. Participants then completed the Big 5 Questionnaire (John, Naumann, & Soto 2008), followed by the Ryff Purpose subscale (Ryff & Keyes 1995), and finally a questionnaire on learning styles. The learning styles questions were only included to mask which scales we were genuinely interested in, and we did not analyze the results by learning style.

After completing the scales participants were presented with an instruction screen where they were told they were about to view a profile of a professor at Cornell who is up for tenure and that “The university is seeking student input on the decision of whether or not to grant tenure to these individuals. You will be randomly assigned to give input about one of these professors. Ultimately some professors will be granted tenure and some will not. Your input is important.” Participants were also told that the three factors on which professors should be evaluated are Research, Service, and Teaching, and definitions of each were provided.

THE DOLEZAL EFFECT

Experimental Manipulation. Participants were randomly assigned to see a photo paired with text featuring a man in professional attire who was either White or Black, and young or middle aged. Participants could see one of 4 possible photos. The accompanying text was always the same: Professor Green was described as having outstanding research and service, but atrocious ratings on student course evaluations.

Dependent Measures. Students then indicated on a seven-point Likert scale how likely they would be to vote to give professor green tenure and on a following page filled in a text entry to explain why they made their decision. After that participants filled out a demographics questionnaire (gender, age, ethnicity, sexual orientation household income, year at Cornell, and undergraduate college at Cornell), then were asked what they believed the purpose of the study was. Finally, they were asked a memory check questions, to verify that they had indeed noticed our experimental manipulation (professor race) and that they had remembered how he had been characterized across the three dimensions (research, service and teaching) they were told to consider. Lastly, participants were asked whether they had believed Professor Green to be a real professor and were debriefed.

Results:

Study 1

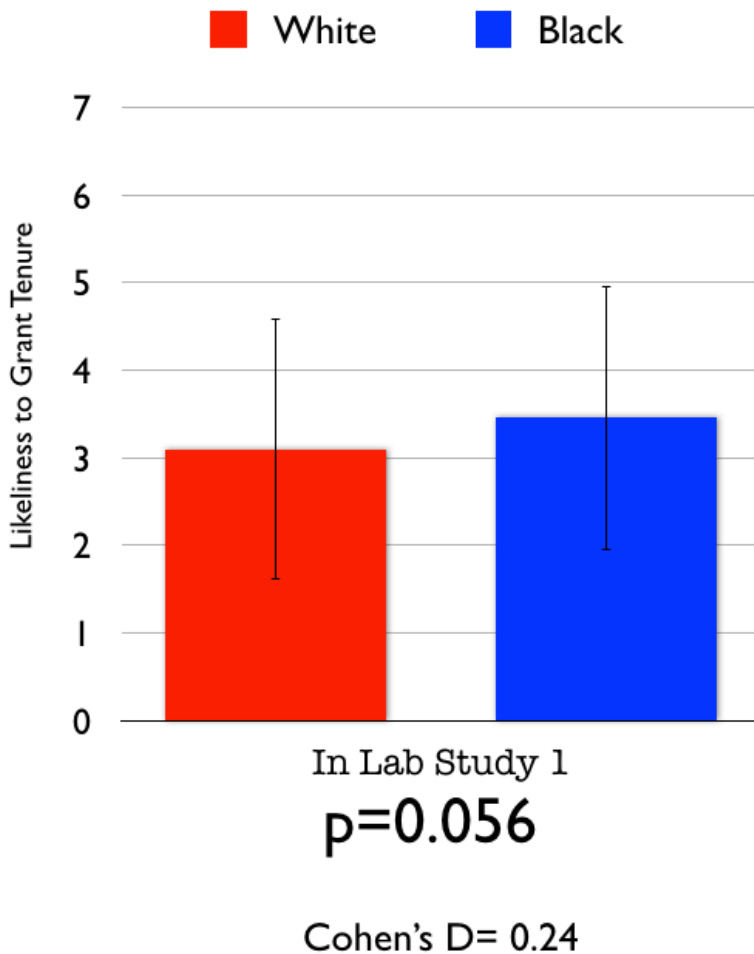


Figure 1: Profiles with Black photos were more likely to receive tenure $p=0.056$, Cohen's $D=0.24$ Error bars are cell standard deviations.

Professor age had no impact and is not discussed further. Results for the young and middle-aged photos of each race are combined for the rest of these analyses. The overall rate of granting tenure across all dimensions was 3.28 on a 7-point Likert scale indicating that the profile was, as intended, a borderline decision.

Profiles of professors paired with Black photos were significantly more likely to receive tenure $p=0.056$, (though even with this elevation to an average Likert scale score of 3.47 this does

THE DOLEZAL EFFECT

remain below the midpoint 4 on the scale). Specifically, the two Black professor profile photos averaged 3.47 on a 7 point Likert scale of likeliness to grant tenure (where 1 was very unlikely and 7 very likely) and the two White professor profile photos averaged a 3.09 on that 7 point Likert scale.

The Big 5 dimension of Openness was a significant moderator $p < 0.05$ ($p < .026$). The Big 5 dimension of Agreeableness was significant at $p = 0.045$.¹

Sense of purpose in life, and the Big 5 dimensions of Extraversion, Conscientiousness and Neuroticism, had no significant main effects. Participant race had no main effect.

Study 2

Participants:

Participants for Study 2 consisted of 222 Cornell undergraduates recruited through SONA who participated for the compensation of course credit. Approximately 83% of participants (184) were female. Participants were drawn from all 7 of Cornell's undergraduate colleges; however, 47% were from the College of Human Ecology, which may explain the greater gender imbalance than observed in Study 1. Of the participants, 112 were White/Caucasian, 61 were Asian/Asian-American, 10 were Black/African-American, 8 were Latino, 17 were multiracial, and 14 identified as other or declined to state. Participants ages ranged from 18 to 23 with over 90% between 18 and 22, and represented a diverse array of family financial back grounds with incomes ranging from between \$20,000 and 50,000 to more than \$200,000, with more than \$200,00 being the modal house hold income value, and the median being between \$170,000 and \$200,000.

¹ These analyses and values reported are found when assessing purpose as a continuous measure, if instead purpose is analyzed as high and low categories in a median split, the big 5 Dimension of Agreeableness is just a trend, Openness remains significant at $p < 0.05$ and the impact of professor race is just barely significant at $p < 0.05$, with $p = 0.04956$. Additionally when analyzed this way this also causes participant race to have a weekly trending interaction with professor race and a weakly trending three way interaction with professor race and sense of purpose

Methods:

Design:

The current study employed a 3 (Stimuli Race: Black vs. White vs. no race control) X 2 (Stimuli Purpose: purpose writing prompt vs control) design.

Procedure:

In contrast to study 1, study 2 was completed online. Upon selecting the study participants were first shown an informed consent document on their screen. Again, after providing their informed consent they were told that we were seeking their input on “a real-life situation” and that we would also be collecting info on their “personality types and learnings styles”. Participants completed the Big 5 Questionnaire (John, Naumann, & Soto 2008), the other two scales present in study 1 were not included in study 2.

Experimental Manipulation 1.

After completing the Big 5 participants were presented with a writing prompt instruction where they either were asked to write about their sense of purpose in the purpose condition or to write about the last movie they saw in the control condition.

Experimental Manipulation 2.

After completing the writing task were participants presented with an instruction screen where they were told they were about to view a profile of a professor at Cornell who is up for tenure and that “The university is seeking student input on the decision of whether or not to grant tenure to these individuals. You will be randomly assigned to give input about one of these professors. Ultimately some professors will be granted tenure and some will not. Your input is important.” Participants were also told that the three factors on which professors should be evaluated are Research, Service, and Teaching, and definitions of each were provided.

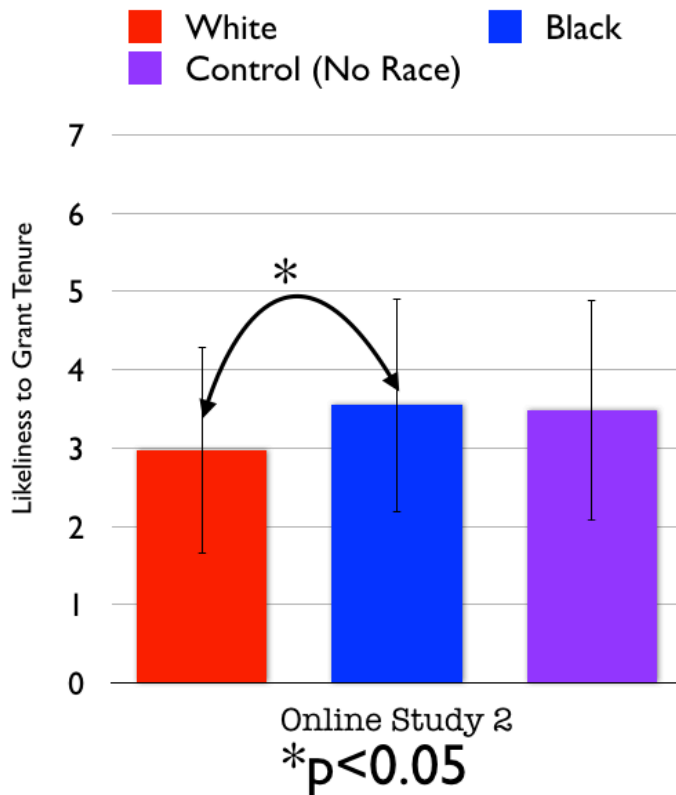
THE DOLEZAL EFFECT

Participants were randomly assigned to see a photo (or no photo in the control condition) paired with text featuring a young man in professional attire. Participants could see one of 3 possible screens, text with no photo, text with a photo of a White man, or text with a photo of a Black man. The accompanying text was always the same as in study 1: Professor Green was described as having outstanding research and service, but atrocious ratings on student course evaluations.

Dependent Measures. Students then indicated on a seven-point Likert scale how likely they would be to vote to give Professor Green tenure and on a following page filled in a text entry to explain why they made their decision. After that participants filled out a demographics questionnaire (gender, age, ethnicity, sexual orientation household income, year at Cornell, and undergraduate college at Cornell), then were asked what they believed the purpose of the study was. Finally, they were asked a memory check questions, to verify that they had indeed noticed our experimental manipulation (professor race) and that they had remembered how he had been characterized across the three dimensions (research, service and teaching) they were told to consider. Lastly, participants were asked whether they had believed Professor Green to be a real professor then read a debrief statement.

Results:

Study 2



ANOVA $p=0.034$; Cohen's $D=0.43$

Estimated Marginal Mean Difference White vs Black $p=0.047$;

Estimated Marginal Mean Difference White vs Control $p=0.12$

Figure 2: There was a significant effect of professor race $p<0.05$. According to estimated marginal mean differences profiles with white photos were significantly less likely to receive tenure than profiles with Black photos $p<0.05$ ($p=0.047$) and were trending to be less likely to receive tenure than profiles with no photo $p=0.12$, Cohen's $D=0.43$. There was no significant or trending difference between the black profile condition and the no photo profile condition. Error bars are standard deviations.

In study 2 professor profiles paired with White photos were significantly less likely to receive tenure than either professor profiles paired with Black photos or profiles paired with no

THE DOLEZAL EFFECT

photo at all $p < 0.05$ ($p = 0.034$). In follow up estimated marginal mean differences there was no significant difference between profiles paired with Black photos and the no photo control condition, but a profile paired with the White photo trended towards differing from the control condition $p = 0.12$, and differed significantly from the Black photo pairing condition $p < 0.05$ ($p = 0.047$). The Black photo profiles averaged a 3.55 likeliness to grant tenure on a seven-point Likert scale of likeliness to grant tenure (where 1 was very unlikely and 7 very likely), the no race control condition averaged 3.49, and the White photo profile averaged a 2.97 on that seven-point Likert Scale.

The Big 5 personality dimension of Conscientiousness had a trending main effect, $p = 0.075$. None of the other Big 5 personality dimensions were significant or trending. Sense of purpose in life had no main effect, and there were no significant or trending interactions for any variables, notably these null findings include participant race, although the participant race main effect $p = 0.212$ and the interaction of participant race and professor race $p = 0.259$ do not necessarily rule out the possibility that there may be some effect or a weak effect of participant race.

Study 1 & 2 Combined Meta-Analysis Results:

Following the procedure and advice set forth by Goh et al (2016) a mini-meta analysis was run on the race main effect of Study 1 and 2 based on the estimated marginal means. This meta-analysis combining the results of study 1 and 2 indicates a significant difference between white and black professor conditions $p < 0.01$ ($p = 0.0036$).

Study 1 & 2 Discussion:

Study 2 replicates the key findings from Study 1. Combined both studies indicate Whites were significantly less likely to receive tenure. In both studies Whites were significantly less likely

THE DOLEZAL EFFECT

to receive tenure. Since study 1 was carried out in the lab and some of our research assistants are ethnic minorities, the possibility of experimenter demand characteristics could not be discounted. The replication in study 2 despite the remoteness of the experimenter, and greater anonymity of participants when they can participate from their dorm rooms or campus libraries indicates that experimental demand characteristics are less likely.

The role of Big 5 personality dimensions after two studies is unclear. Despite Openness being significant in Study 1, it was not in Study 2. No personality dimension was significant or trending in both studies. Therefore, personality dimensions are unlikely to be a major factor. Despite our hope that purpose might buffer against racial biases and discomfort as it had in Burrow and Hill (2013) our purpose manipulation in Study 2 did not produce any noticeable effects. Thus, the weakly trending interaction effect of purpose in footnote 1 seems quite possibly attributable to random chance.

Study 1 and 2 both successfully zeroed in on the grey area where racial bias is most likely to have an effect (Dovidio & Gaertner 2000). In the aggregate in both studies willingness to confer tenure remained less than a single point on the scale away from indecision. It is however, notable that in both Study 1 and 2 it was below this midpoint for all conditions. Students may have taken the bad teaching evaluations to heart as a domain which directly affects their experience. Thus, rather than characterizing this effect as a boost to minorities, or a detriment to Whites, it might be better characterized as unwillingness to forgive Whites. What is particularly remarkable about study 2, is the difference between the White condition and the no race condition. Somehow being visually primed with a White profile, made participants less forgiving of the professor's supposed bad teaching.

After two studies involving college student participants samples, and a university bounded question like professor tenure, our next step was to test whether this could be generalized to

promotion decisions in other contexts and a general population sample. Also, the results of study 2 in particular, and differential forgiveness for violating a social script (by being a bad teacher), lead us to wonder what would happen if the deviation, instead of being one of objective quality as in study 1 and study 2, was instead a deviation from racial cultural norms.

Study 3

Participants:

Participants for Study 3 consisted of 210 general population adults recruited through amazon mechanical turk who participated for \$0.50. Approximately 45% of participants (94) were female, 55% were male (113) and 3 participants declined to state their gender. Of the participants, 154 were White/Caucasian, 16 were Asian/Asian-American, 14 were Latino, 11 were Black/African-American, 1 was Native American, 9 were multiracial, and 2 identified as other or declined to state. Participants ages ranged from 18 to 64. While participants again represented a diverse array of family financial backgrounds, this sample was more similar to the general population and although incomes ranged from less than \$20,000 to more than \$200,000, the modal and median income bracket was \$20,001 and \$50,000.

Methods:

Design:

The current study employed a 3 (Stimuli Race: Black vs. White vs. Asian) X 2 (Stimuli racially stereotypical or atypical music club: Classical Music vs. Hip-Hop) design.

Procedure:

Like study 2, study 3 was completed online, but unlike study 2, participants selected the study through amazon mechanical turk instead of the university system. Upon selecting the study participants were first shown an informed consent document. Participants then completed

THE DOLEZAL EFFECT

the Big 5 Questionnaire (John, Naumann, & Soto 2008), followed by the Ryff Purpose subscale (Ryff & Keyes 1995), and finally a questionnaire on learning styles. The learning styles questions were only included to mask which scales we were genuinely interested in, and we did not analyze the results by learning style.

Participants then read about a fictional Everdale School District. In order to mirror the same dynamic of tenure or fire decisions from studies 1 and 2, participants were told that in Everdale tenure is conferred after two years, but that this year the school district is facing budget cuts and some second-year teachers will have to be fired. Participants were told they would be randomly assigned to see one teacher profile and were instructed to evaluate it on the basis of student performance, service, and student teacher approval. Before moving on to the experimental manipulation participants completed an attention check where they are asked after what year are tenure decisions made in Everdale.

Experimental Manipulations.

After completing the attention check participants were randomly assigned to see a photo paired with text featuring a middle-aged man in professional attire. Participants could see one of 3 possible photos, a middle-aged White man, a middle-aged Black Man, or a middle-aged Asian man. The White and Black photos, were the same middle-aged photos from study 1, but differed from study 2 which had used the young photos. This was done to disabuse the notion that the results of studies 1 and 2 could be attributable to something uniquely bizarre about the young photos. An Asian condition was also added to see whether participants treated a model minority like White profiles or more like Black profiles (Wong, et al. 1998). The name for the teacher in the text was always Mr. Lee, since Lee is a common surname for individuals of Asian ancestry and also occasionally found as a White or Black last name, specifically according to the 2010 census 42% of Lee's in the US identified as Asian, 36% identified as White, and 16% identified

THE DOLEZAL EFFECT

as Black (Comenetz 2016). The middle-aged Asian photo was selected from the same stock image source as the other faces and chosen for similarity of age, attire, camera angle and general aesthetic, but was not pretested for age and attractiveness.

The accompanying text featured one of two conditions which changed the description of how Mr. Lee fulfilled his service obligations. Participants always read that student performance was slightly above average and student approval was slightly below average. The third category, service, stated “Mr. Lee’s primary service involvement is that he serves as the faculty advisor for a student club called the Everdale [Hip Hop/Classical Music] Appreciation Society. This was chosen because classical music is a stereotypically White interest, while Hip Hop is a stereotypically Black interest.

Dependent Measures. Students then indicated in a dichotomous choice whether they would fire Mr. Lee or grant him tenure. They were also asked on sliding scales how certain they were about their decision, how in favor of granting tenure they were, and how strongly against it they were. On the following page participants filled in a text entry to explain why they made their decision. After that participants filled out a demographics questionnaire (gender, age, ethnicity, sexual orientation household income, and state of residence), then were asked what they believed the purpose of the study was. Finally, they were asked a memory check questions, to verify that they had indeed noticed our experimental manipulations (teacher race, and what club Mr. Lee supervised) and that they had remembered how he had been characterized across the other two dimensions they were told to consider (student performance, and student evaluations). Lastly, they finished the survey and received instructions on how to receive their compensation.

Results:

THE DOLEZAL EFFECT

When analyzing the entire sample looking at the binary tenure/fire decision variable and running a binary logistic regression nothing is significant. The Big 5 dimension of neuroticism is trending ($p=0.077$) and openness is weakly trending ($p=0.122$). For the continuous slider scale dependent variable measure of how in favor participants were of granting tenure, when analyzing the entire participant sample, the only finding which is significant is the Big 5 dimension of openness $p<0.05$ ($p=0.018$). The main effect of which club the teacher supervises (classical music vs hip hop) was trending ($p=0.098$), and most importantly so is both a main effect of participant race ($p=0.067$) and a three-way interaction of teacher race, club supervision, and participant race ($p=0.067$). Since the $N=210$ sample consists of only 52 non-White participants (a little less than a quarter of the sample) and participant race appears to possibly play a role the analyses were all re-run to specifically examine the White portion of the sample.

Study 3: Dolezal Effect

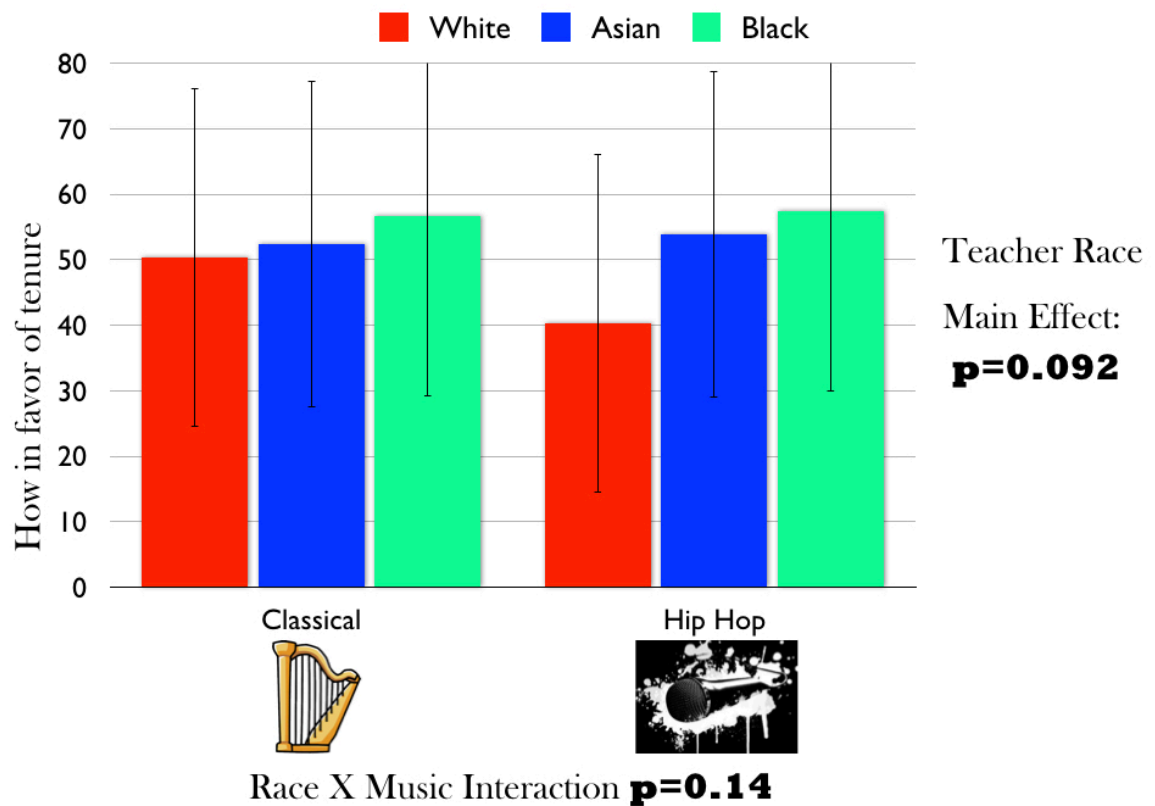


Figure 3: There was a trending effect of teacher race $p=0.092$, and a trending music race interaction of 0.14.

According to estimated marginal means differences profiles with white photos were less likely to receive tenure than profiles with Black photos, but there was no significant or trending difference between the black profile condition and the Asian photo profile condition. Error bars are standard deviations.

Rerunning the analysis on the white participants, for the dichotomous tenure/fire dependent variable little changes. Neuroticism becomes significant $p<0.05$ ($p=0.033$), and nothing trends. But some changes are observed for the continuous variable assessing how in favor participants are of granting the teacher tenure. The Big 5 dimension of openness is significant $p<0.05$ ($p=0.010$). And there is a trend for teacher race ($p=0.092$). Specifically estimated

THE DOLEZAL EFFECT

marginal mean differences reveal that the white teacher condition is trending to be less likely to be tenured than the black teacher condition ($p=0.095$) and there is no apparent difference between the Asian condition and the Black condition.

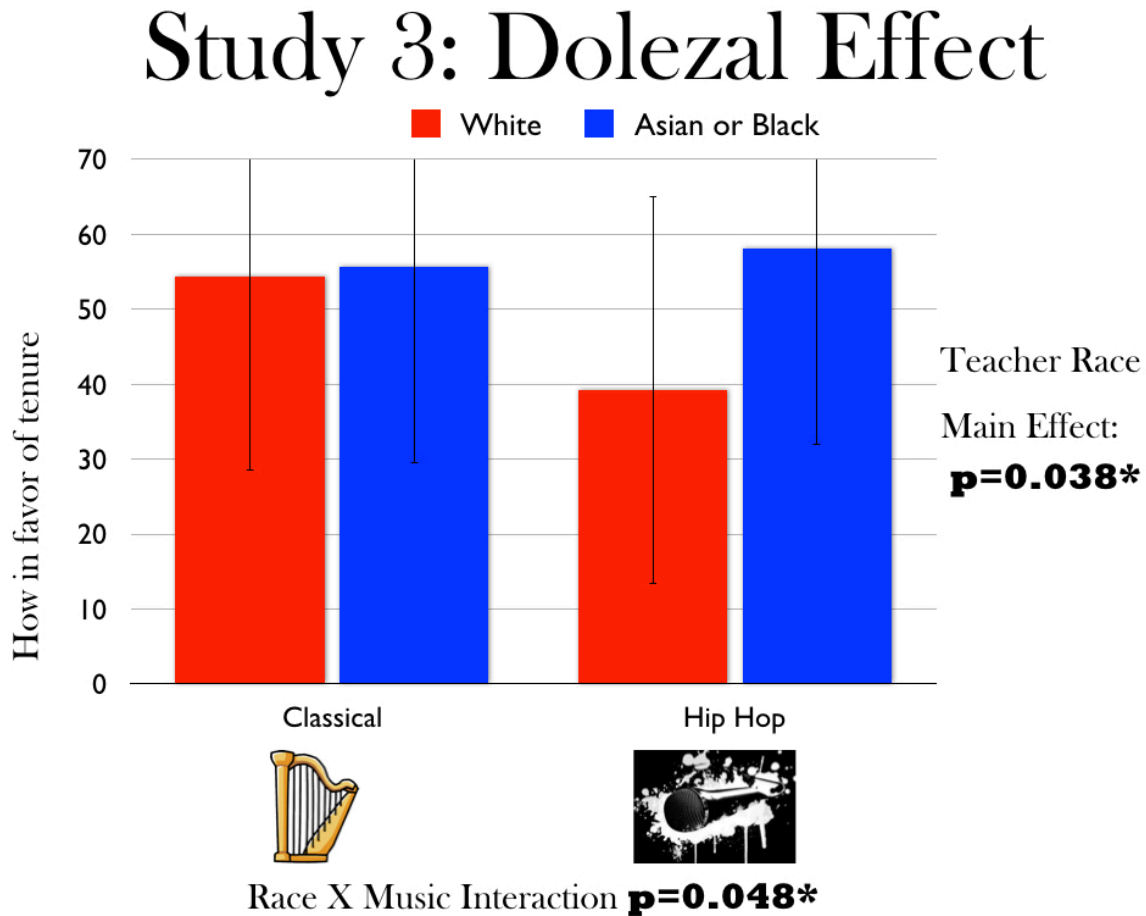


Figure 4: There is a significant interaction, such that White teachers who supervise the hip-hop club are significantly less likely to receive tenure ($p<0.05$). Additionally, the main effect of teacher race is (White vs Asian or Black) is also significant ($p<0.05$). Error bars represent standard deviation.

Because the Asian photo conditions behaved indistinguishably from the Black photo conditions and two are combined for the purposes of the following analyses.

For the categorical two option dependent variable measure to the question “based on the description you have just read, would you grant Mr. Lee tenure or fire him?” there was a strongly

THE DOLEZAL EFFECT

trending nearly significant interaction of music condition and professor race $p=0.053$ when analyzed using a chi square test. For this categorical dependent variable, no other variables were significant, though there was a moderate trend ($p=0.086$) for the Big 5 Neuroticism dimension, and a weak trend ($p=0.137$) for the Big 5 Openness dimension. Specifically, in the condition of the White Teacher who supervised the classical music club 54.3% were in favor of granting tenure, whereas only 37.1% were in favor of granting when the white teacher supervised the hip hop club, 51.5% were in favor of the granting tenure to Black and Asian teachers who supervised the classical music club, and 60.6% were in favor of granting tenure to the Black and Asian teachers who supervised the hip hop music club.

The sliding scale dependent variable for the question how in favor are you of granting tenure to Mr. Lee featured a significant interaction ($p<0.05$). Specifically, White teachers who supervised the hip hop club (the racial stereotype incongruent club) were significantly less likely to receive tenure than White teachers who supervised a classical music club, or Black and Asian teachers who supervised any club. We also see a significant main effect of teacher race $p=0.038$ ($p<0.05$), replicating study 1 and 2 here; however, this main effect appears to be driven by the interaction effect, where White teachers in violation of the White stereotype were least likely to be granted tenure. On the sliding scale (0 to 100 where 0 was not at all in favor and 100 was completely in favor) the white teacher hip hop club supervisor condition the mean response was 39.2, whereas in the white teacher classical music club supervisor condition the mean response on the sliding scale was 54.3, the Black and Asian teachers in the hip hop condition averaged 58.1 on the sliding scale, and the Black and Asian teachers in the classical music condition averaged 55.6 on the sliding scale.

The sliding scale dependent variable question of how against granting tenure are you, had no significant moderators and is not discussed further. For the sliding scale question about how certain participants were about their decision to grant tenure to or fire Mr. Lee, the only significant moderator was music club condition, where participants were marginally more certain about their decision when

THE DOLEZAL EFFECT

Mr. Lee supervised the classical music club. The interaction of teacher race and decision was also a trending moderator $p=0.074$ (trending toward being least certain about firing Black and Asian teachers).

Discussion/Conclusions

Participants appeared to punish White teachers who supervised the stereotypically incongruent Hip Hop club. We are calling this apparent aversion to perceived cultural appropriation the Dolezal Effect, after the public backlash to Rachel Dolezal for what her detractors perceived as cultural appropriation, and an impermissible encroachment on black identity by a White woman. We believe a similar aversion is at play here. In studies 1 and 2 Whites were punished for deviating from the culturally accepted stereotype of competence, in study 3, while we also replicated the pro-minority main effect, this was driven by an aversion to the White teacher supervising a Hip-Hop club, the Dolezal effect interaction. A small distinction from study 1 and 2, where all conditions resulted in evaluations below the midpoint on the scale, here both sides of the scale were in play with the Black and Asian condition averaging over 50% in favor of bestowing tenure, while the White teacher who supervises hip hop was below the 50% mark.

While we had hoped that perhaps sense of purpose in life might show so debiasing utility as a trait variable if not a manipulated state variable, unfortunately as in study 2, study 3 did not find an effect of purpose. Similarly, no consistent story for any of the big 5 personality dimensions emerged either.

Our results and the Dolezal interaction effect show that context is important. For high status jobs like teachers or professors, when ethnic minority employees have clearly demonstrated simply by being in the room under consideration that they are not congruent with their negative stereotype then they may receive a boost their White colleagues do not just as in Axt Ebersol and Nosek (2016) and Williams and Ceci (2015) particularly if the White colleague (despite equal matching) is in some way incongruent with the White positive stereotype. At least when the person is perceived as at minimum

THE DOLEZAL EFFECT

potentially qualified (Ceci & Williams 2015; Mendes and Koslov 2013). It is also important to remember that a formal consideration of promotion is an occasion that allows for a high amount of cognitive resources to be devoted to that task, cognitive resources that Mendes and Koslov (2013) make clear are prerequisite conditions for a pro-minority/anti-majority bias to be displayed.



Figure 5: If success is envisioned as a ladder the ladder minorities and stigmatized groups have to climb could be construed as one covered in thorns, but any particular rung of the ladder may be thornier or smoother than the equivalent rung on the white ladder.

As shown in the metaphorical model of figure 5 our results need not undermine the predominant findings of systemic inequalities in the literature, but rather add some nuance to this picture. Promotion decisions for high status jobs are just one of the many rungs on the ladder to success, and if the character of this rung is that of a harder path for whites that does not mean that we don't still see an overall advantage for Whites in society when all the rungs on the ladder are taken together. Understanding when and how rungs on the latter can deviate from the typical pattern of greater hardship and

disadvantage for minorities is important for understanding the big picture and may be instrumental in learning how to smooth out other rungs and combat bias at other steps.

Future Directions/Limitations

Promote vs. Fire vs. Promote vs. Status Quo

One feature of tenure decisions that made them a particularly logical extension to the existing literature on hiring decisions, is the galvanizing nature of the decision, where on one hand there is the decision to confer tenure whereas on the other hand there is essentially the decision to fire. This sharp divide closely mirrors the nature of hiring decisions and the existential separation between hiring someone on the one hand vs keeping them outside of the company/school/organization on the other. However, most real-world promotion decisions do not have the same immediacy of exit, and are typically a choice between promotion on the one hand and preservation of the status quo on the other at least in the short run. Therefore, one future direction is to investigate whether/how a promote vs status quo decision differs from hiring decisions and promote vs fire decisions.

Real Decision Makers

Study 1 and Study 2 used a sample drawn from the Cornell University undergraduate student body. Study 3 used a sample from m-turk. Though college students certainly understand what it means to be promoted or fired, they are not by in large themselves very experienced with being the person responsible for making decisions about who to retain. Similarly, though an m-turk sample is perhaps more representative of the general population than a student sample, and is certainly more age diverse, it still remains the case that most people's occupational responsibilities do not include being the one to make executive decisions about retention and promotion. Therefore, replicating the results of these studies in a sample with real decision makers is an important direction for future work to consider.

Relatedly while we presume that the effect found in study 3 is cut from the same cloth as the findings of

THE DOLEZAL EFFECT

study 1 and 2 it is important to remember that these were two different population samples (undergrads in studies 1 and 2, and m-turkers in study 3), and it remains possible that different things could be driving these effects for the different populations.

Ally or Cultural Appropriator?

Perhaps the most consequential question for a Dolezal effect, is what does it mean for allies? Does the same psychological mechanism which causes people to feel unease about cultural appropriation also create a barrier for those who seek to help as allies? An aversion to cultural appropriation could be seen as an aversion to a specific type of stereotype inconsistency. Where the bounds of that ends is an important question. Study 3 suggests that some objectively neutral criteria, like supporting a music club, receives a negative valence for whites when it is inconsistent with their racial stereotype. Does this mean that would be allies are discouraged from entering helping fields like civil rights law, immigration law, or social services? And if so does such discouragement contribute to a ghettoizing and perpetuate racializing of social strata?

All Things Being Equal, But Are Opportunities Really Equal?

Across these three studies where either a Dolezal interaction effect or pro-minority main effect was found as well as in the prior work of Axt Ebersol and Nosek (2016), Williams and Ceci (2015), and Mendes and Koslov (2013) these effects were found when qualifications were held constant. Ceci & Williams (2015) makes clear that this does not extend to situations where one candidate is clearly superior to another. But the matter of how likely are any two prospective candidates for promotion (or for hiring) are to have equal resumes remains unresolved. In fact, it could be objectively wise to value an equally matched minority applicant more highly if they face greater adversity in achieving each of those same resume lines. Mendes and Koslov (2013) indicate that the pro-minority biases may be constrained to situations where participants do not

THE DOLEZAL EFFECT

suffer from increased threat or cognitive load. Whom senior figures choose to rely on or delegate important tasks too when under pressure may result in greater opportunities in the work place for majority group members. Also the fact that some of the studies (Mendes and Koslov 2013; Mizel 2018) finding pro-minority biases, found this to be greatest among individuals with greatest anti-Black biased IAT scores raises the question of how these minority employees are treated on a day in day out biases outside of these focused evaluative moments, and whether those same employees who reap benefits may also be at greater risk for frequency of suffering micro-aggressions (Sue, Capodilupo, & Holder 2008; Sue et al 2007; etc.). This taken along with Richeson and Trawalter's (2008) finding that desire not to be prejudiced disrupts White persons' attentional processes, suggest that the same individuals who may in formal consideration settings like promotion decision offer a bonus to underrepresented minorities, may, in day to day interactions, provide them with less opportunities and interact with less friendliness and more agitation.

A Rose by Any Other Name?

Conceptual replication across context and medium is an integral part of the scientific process. These three studies relied on profile photos in order to manipulate race and find the resulting pro-minority main effect and the anti-appropriation Dolezal effect in study 3. Future work should investigate if these biases in promotion decisions persist when using alternative methods to manipulate employee race, such as manipulation of names. Until such future work is conducted we cannot definitively rule out the possibility that our results could be an artifact of our study stimuli and/or the particular attractiveness/unattractiveness of our photo stimuli.

These three studies investigated contrasts between treatment of White and Black employees facing promotion or termination decisions. While study 2 also included a no race and Study 3 included an Asian condition, there remains much to be explored both in terms of

THE DOLEZAL EFFECT

raceless default norms, and the treatment of Asian employees as well as every other race not included in this study, such as the treatment of Latino employees. The world is full of biases; however, it appears that when it comes to promotion decisions three studies show these biases can run counter to the traditional direction. The Dolezal effect found in study 3 shows that when stereotypes are confounded positively stereotyped groups can face a negative backlash for their deviation/straying into cultural appropriation. Future work must explore what this means for would be allies, as well as where and by what mechanism underrepresented groups fall off the pipeline and into underrepresentation if not at the point of formal promotion decisions.

REFERENCES

- Axt, J. R., Ebersole, C. R., & Nosek, B. A. (2016). An unintentional, robust, and replicable pro-Black bias in social judgment. *social cognition*, 34(1), 1-39.
- Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American economic review*, 94(4), 991-1013.
- Burch, T. (2015). Skin Color and the Criminal Justice System: Beyond Black-White Disparities in Sentencing. *Journal of Empirical Legal Studies*, 12(3), 395-420.
- Burrow, A. L., & Hill, P. L. (2013). Derailed by diversity? Purpose buffers the relationship between ethnic composition on trains and passenger negative mood. *Personality and Social Psychology Bulletin*, 39(12), 1610-1619.
- Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological science*, 15(10), 674-679.
- Ceci, S. J., & Williams, W. M. (2015). Women have substantial advantage in STEM faculty hiring, except when competing against more-accomplished men. *Frontiers in psychology*, 6, 1532.
- Comenetz, J. U.S. Census Bureau (2016). Frequently Occurring Surnames in the 2010 Census. Retrieved from <https://www2.census.gov/topics/genealogy/2010surnames/surnames.pdf>
- Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of personality and social psychology*, 83(6), 1314.

THE DOLEZAL EFFECT

- Dinovitzer, R., Garth, B. G., Plickert, G., Sandefur, R., Sterling, J., Wilkins, D., & Nelson, R. L. (2014). After the JD III: Third results from a national study of legal careers. American Bar Foundation and NALP Foundation for Law Career Research and Education.
- Dovidio, J. F., & Gaertner, S. L. (2000). Aversive racism and selection decisions: 1989 and 1999. *Psychological science*, 11(4), 315-319.
- Dovidio, J. F., & Gaertner, S. L. (2004). Aversive racism. *Advances in experimental social psychology*, 36, 4-56.
- Garcia-Retamero, R., & López-Zafra, E. (2006). Prejudice against women in male-congenial environments: Perceptions of gender role congruity in leadership. *Sex roles*, 55(1-2), 51-61.
- Glick, P., Gangl, C., Gibb, S., Klumpner, S., & Weinberg, E. (2007). Defensive reactions to masculinity threat: More negative affect toward effeminate (but not masculine) gay men. *Sex Roles*, 57(1-2), 55-59.
- Greenwald, A. G., & Pettigrew, T. F. (2014). With malice toward none and charity for some: Ingroup favoritism enables discrimination. *American Psychologist*, 69, 669-684.
- Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some arguments on why and a primer on how. *Social and Personality Psychology Compass*, 10(10), 535-549.
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm Shift to the Integrative Big-Five Trait Taxonomy: History, Measurement, and Conceptual Issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114-158). New York, NY: Guilford Press.

THE DOLEZAL EFFECT

- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice*, 6(1), 101.
- McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of experimental Social psychology*, 37(5), 435-442.
- Mendes, W. B., & Koslov, K. (2013). Brittle smiles: Positive biases toward stigmatized and outgroup targets. *Journal of Experimental Psychology: General*, 142(3), 923.
- Mizel, M. L. (2018). *A Plea for Justice: Racial Bias in Pretrial Decision Making* (Doctoral dissertation, UCLA).
- Murphy, P. P. (2017, August 14) White Nationalists use Tiki Torches to Light up Charlottesville March. CNN. Retrieved from <https://www.cnn.com/2017/08/12/us/white-nationalists-tiki-torch-march-trnd/index.html>
- Neville, H. A., Lilly, R. L., Duran, G., Lee, R. M., & Browne, L. (2000). Construction and initial validation of the color-blind racial attitudes scale (CoBRAS). *Journal of Counseling Psychology*, 47(1), 59.
- Pager, D., Bonikowski, B., & Western, B. (2009). Discrimination in a low-wage labor market: A field experiment. *American sociological review*, 74(5), 777-799.
- Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of personality and social psychology*, 75(3), 811.
- Quillian, L., Pager, D., Hexel, O., & Midtbøen, A. H. (2017). Meta-analysis of field experiments shows no change in racial discrimination in hiring over time. *Proceedings of the National Academy of Sciences*, 114(41), 10870-10875.

THE DOLEZAL EFFECT

- Ragins, B. R., & Cornwell, J. M. (2001). Pink triangles: antecedents and consequences of perceived workplace discrimination against gay and lesbian employees. *Journal of applied psychology*, 86(6), 1244.
- Richeson, J. A., & Trawalter, S. (2008). The threat of appearing prejudiced and race-based attentional biases. *Psychological Science*, 19(2), 98-102.
- Rosette, A. S., Leonardelli, G. J., & Phillips, K. W. (2008). The White standard: Racial bias in leader categorization. *Journal of Applied Psychology*, 93, 758-77.
- Ryff, C. D., & Keyes, C. L. M. (1995). The structure of psychological well-being revisited. *Journal of personality and social psychology*, 69(4), 719.
- Sue, D. W., Capodilupo, C. M., & Holder, A. (2008). Racial microaggressions in the life experience of Black Americans. *Professional Psychology: Research and Practice*, 39(3), 329.
- Sue, D. W., Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A., Nadal, K. L., & Esquilin, M. (2007). Racial microaggressions in everyday life: Implications for clinical practice. *American psychologist*, 62(4), 271.
- Williams, W. M., & Ceci, S. J. (2015). National hiring experiments reveal 2: 1 faculty preference for women on STEM tenure track. *Proceedings of the National Academy of Sciences*, 201418878. Chicago
- Wong, P., Lai, C. F., Nagasawa, R., & Lin, T. (1998). Asian Americans as a model minority: Self-perceptions and perceptions by other racial groups. *Sociological perspectives*, 41(1), 95-118.
- Zigerell, L. J. (2018). Black and White discrimination in the United States: Evidence from an archive of survey experiment studies. *Research & Politics*, 5(1), 2053168017753862.